Parsing and interpreting wh-in-situ scope

Ming Xiang

The University of Chicago

Parsing and comprehension

- Traditional view: Language comprehension depends on parsing. By parsing, we mean the process to parse an incoming string into structured grammatical representations

- A stronger implicit assumption: There is a direct mapping between parsing and interpretation. The (incremental) output of parsing serves as the (incremental) input of interpretation. Therefore, sentence interpretation should closely track sentence parsing.

Parsing and comprehension

- **Challenges to the traditional view**: There are clear empirical cases where interpretations obtained by comprehenders do not match the possible parses of the sentence
  - lingering garden path ambiguity (Christianson et al. 2001; Qian et al. 2018; Slattery et al., 2013)
  - Local coherence (Tabor, 2004)

(possibly also comparative illusion, Wellwood et al. 2018, NPI grammatical illusion, Xiang et al. 2009)

- *Good enough* comprehension model:

  Interpretations derived through simple (non-structural) heuristics can override interpretations derived from a fully specified grammatical parse (Ferreira et al., 2001, 2002; Christianson et al. 2001; Ferreira & Patson, 2007)

- *Noisy channel* model:

  There is noise/uncertainty over the linguistic input a listener perceives. Listeners may have “edited” the input to their more likely near neighbors (Levy, 2008; Gibson et al. 2013)

The empirical case today:

- We will examine wh-in-situ scope taking, and demonstrate a new case of misalignment between parsing and interpretation
- In particular, we will show that for wh-in-situ sentences with ambiguous scope, the preferred interpretation doesn't necessarily track the preferred parse

The proposal today:

- Our solution relies on a richer model to map the parsing output to interpretations, instead of the more traditional ‘direct mapping’.

- Interpretation will be modeled as the listener’s pragmatic inference about the relevant world states conditioned on an utterance

- Parsing bias does not uniquely determine the interpretation bias, but it will be a built-in component in the pragmatic reasoning model

Potential connections to the intervention effect

- Even the basic wh-in-situ scope taking phenomenon could involve substantial processing complexity
- We need a better understanding of the most common methods in linguistics: acceptability judgments and truth value judgments
- Some of the variability observed in the intervention literature could potentially be accounted for by processing factors

Mandarin wh-in-situ dependencies

约翰	隐瞒了	什么	事？
Yuehan	yingman-le	shenme	shi
John	hide-perf	what	thing

"What did John hide?"

(Huang, 1982; Li, 1992; Aoun & Li 1993; Tsai 1994; Cheng, 1991; 2003)

In Xiang, Wang and Cui, 2015, we provided both comprehension and production evidence to argue for the construction of a covert dependency in real time processing of wh-in-situ :

Comprehension: parsing complexity of wh-in-situ constructions, but not their declarative counterparts, is affected by the intervening CP boundaries

Production: speakers avoid wh-in-situ constructions in order to avoid the long distance covert wh-dependency

How does the parser access the relevant scope position?

$\ldots V_1[_{CP1} \ldots V_2[_{CP2} \ldots WH]$

Emily **announced** her team **discovered** aliens built **which city.**

a. “Emily announced **which city** her team discovered aliens built.”

b. “Emily announced her team discovered **which city** aliens built.”

How does the parser access the scope position?

$$\ldots V_1[_{CP1} \ldots V_2[_{CP2} \ldots WH]$$

Locality bias: There is a strong bias towards the local scope parse.
(Xiang et al., *manuscript in progress*)

Experiment 1: Detecting the locality bias

a.

…find out$[_{CP1}$… know$[_{CP2}$ ... ***WH***]]

a possible dependency

b.

…find out$[_{CP1}$… believe$[_{CP2}$ ... ***WH***]]

an impossible dependency

- Due to the locality bias:
  - there would be more processing cost when the local scope is blocked

    processing cost (b) > (a)
  - The processing cost arises due to the complexity accessing the high scope, or/and complexity in structural reanalysis

c.

. . .**wonder** [$_{\text{CP1}}$. . . know[$_{\text{CP2}}$ ... ***WH***]]

d.

. . .**wonder** [$_{\text{CP1}}$. . . believe[$_{\text{CP2}}$ ... ***WH***]]

In the presence of an anticipatory cue that signals the correct non-local scope position:

- Will it be easier for the in-situ wh-phrase to access the non-local scope position?
- Will we still see the effect of the local scope position?

a.

小王 打听到 工程队 知道 村民们 扩建了 哪座水坝.

Mr. W. **find out** construction-team **know** villagers rebuild **which dam**

“Mr. W. **found out** **which dam** the team **knew** the villagers rebuilt.”

“Mr. W. **found out** the team **knew** **which dam** the villagers rebuilt.”

b.

小王 打听到 工程队 相信 村民们 扩建了 哪座水坝.

Mr. W. **find out** construction-team **believe** villagers rebuild **which dam**

“Mr. W. **found out** **which dam** the team **believed** the villagers rebuilt.”

c.

小王 想弄明白 工程队 知道 村民们 扩建了 哪座水坝.

Mr. W. **want-to-figure-out** construction-team **know** villagers rebuild **which dam**

“Mr. W. **wondered** **which dam** the team **knew** the villagers rebuilt.”

d.

小王 想弄明白 工程队 相信 村民们 扩建了 哪座水坝.

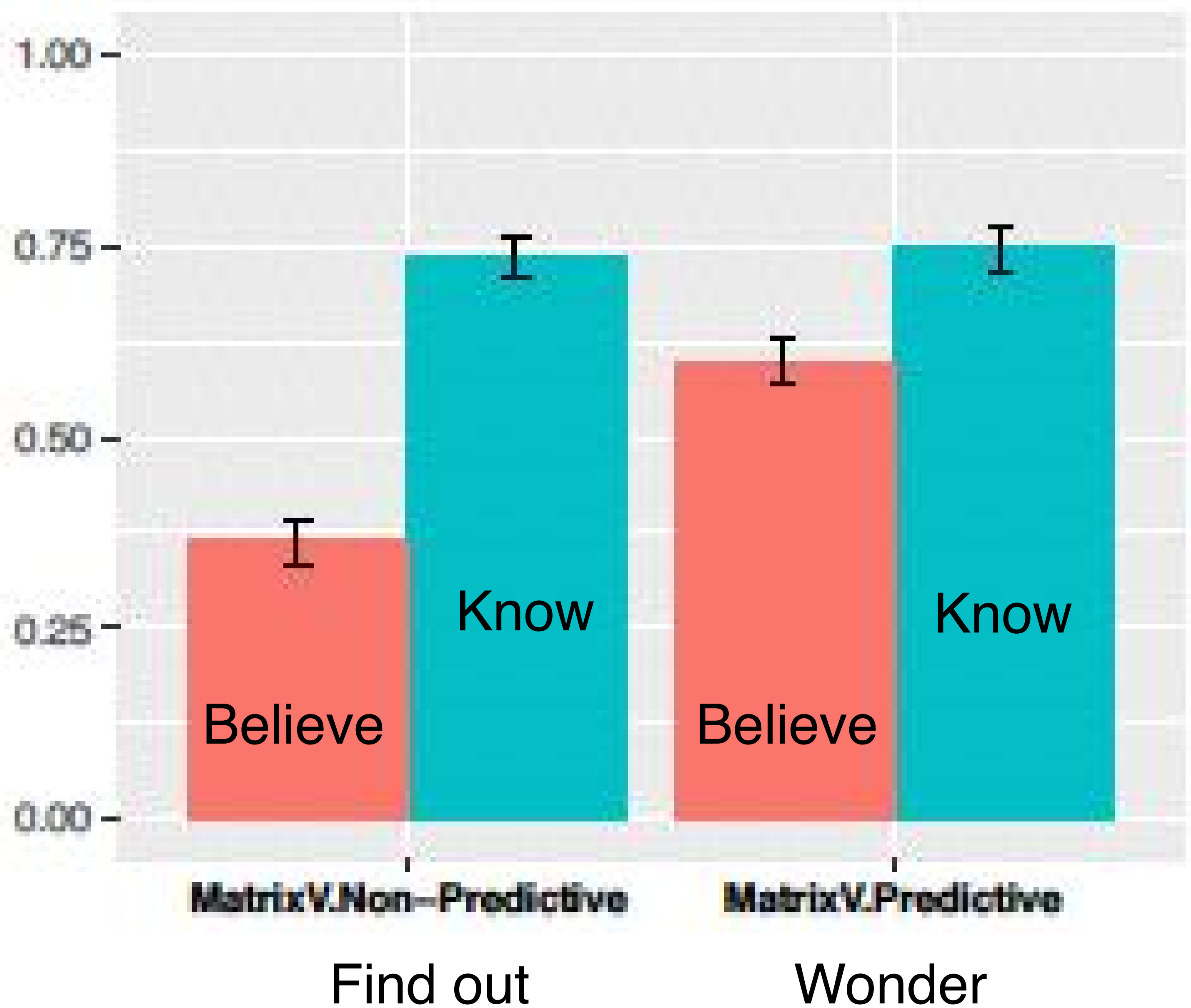
Mr. W. **want-to-figure-out** construction-team **believe** villagers rebuild **which dam**

“Mr. W. **wondered** **which dam** the team **believed** the villagers rebuilt.”

Experiment 1 procedure

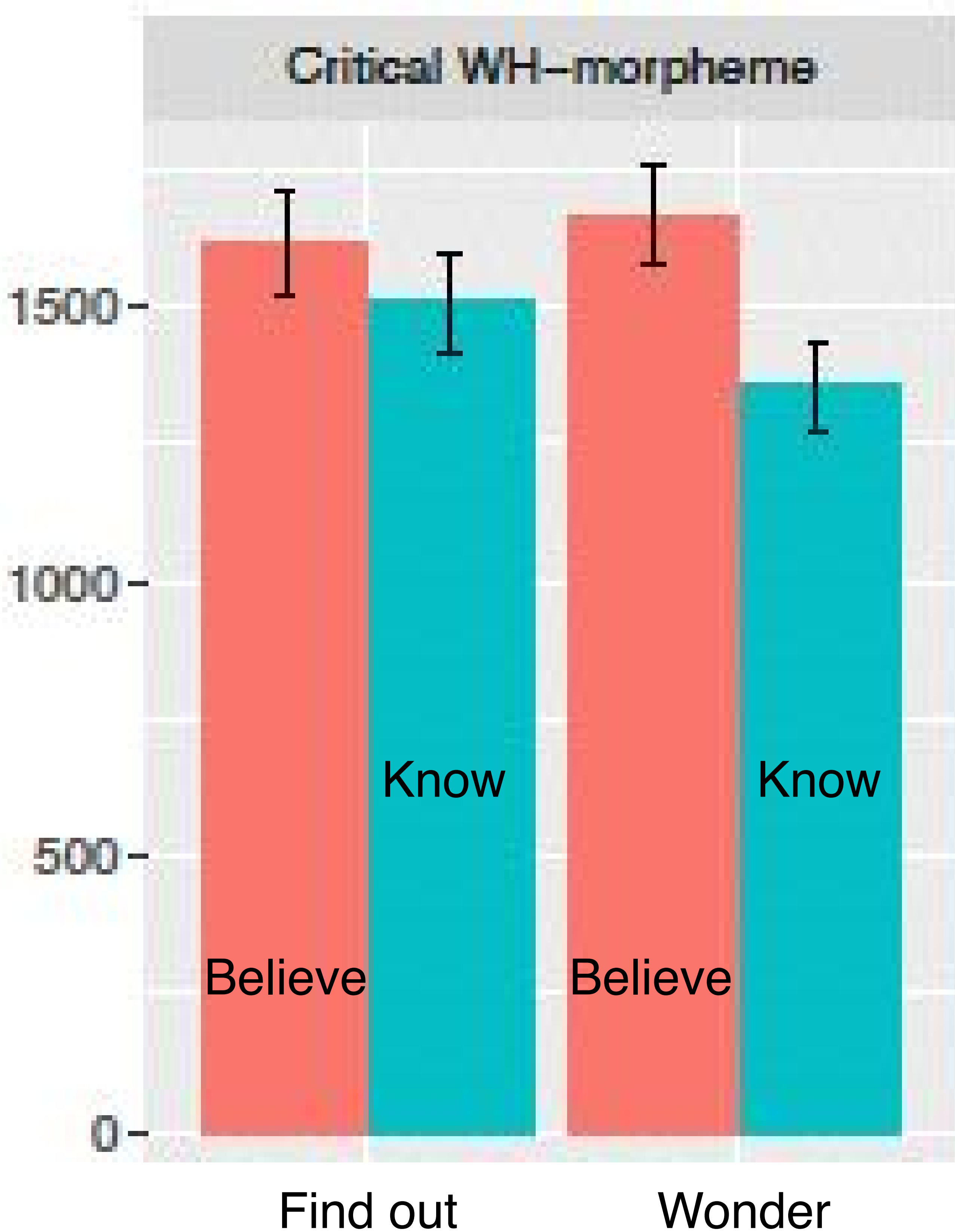
- Eye-tracking reading
- Yes-no acceptability judgment task after each trial
- Critical word (CW) is the sentence final wh-phrase
- 40 items
- 40 subjects

Acceptability rating results



- Main effect of Locality
- an interaction between locality and matrix verb: an early predictive cue (e.g. *wonder*) does not completely overcome the locality bias

Regression reading time at the wh-morpheme



- Main effect of Locality
- An early predictive cue (e.g. *wonder*) does not have any effect on the locality bias

- Summary of the findings:
    - There is a strong local scope bias. When the local scope is blocked, it is difficult for the parser to carry out the reanalysis that arrives at the non-local scope parse
    - The locality bias is even present when there is an early predictive cue and the only grammatical parse is the non-local scope

Experiment 2:

a.

…find out[**CP1**… know[**CP2** ... ***WH***]]

b.

longer regression
RT time

…find out[**CP1**… believe[**CP2** ... ***WH***]]

lower acceptability rating

a.

…find out[$_{CP1}$… know[$_{CP2}$ ... ***WH***]] **ambiguous**

b.

longer regression RT time

…find out[$_{CP1}$… believe[$_{CP2,\ -Q}$ ... ***WH***]]

unambiguous

lower acceptability rating

The locality bias account: when the local scope is blocked, **reanalysis** from a local parse to a non-local parse is costly

An alternative account: The ambiguous condition is independently easier than the unambiguous one (e.g. ***the ambiguity advantage***, Traxler, Pickering, and Clifton, 1998; Logacev and Vasishth, 2015; Swets et al. 2008)

Two additional conditions:

c.

…know [$_{\text{CP1}}$…find out[$_{\text{CP2}}$ ... ***WH***]] **ambiguous**

d.

…believe [$_{\text{CP1}}$ …find out[$_{\text{CP2}}$ ... ***WH***]]

unambiguous

Predictions for Experiment 2:

The locality bias account:

no difference in acceptability and reading time between the two new conditions, and both should be relatively easy to process

The ambiguity advantage account:

Acceptability: ambiguous condition > unambiguous

RT: ambiguous condition < unambiguous

a.

小王 打听到 工程队 知道 村民们 扩建了 哪座水坝.

Mr. W. **find out** construction-team **know** villagers rebuild **which dam**

b.

小王 打听到 工程队 相信 村民们 扩建了 哪座水坝.

Mr. W. **find out** construction-team **believe** villagers rebuild **which dam**

c.

小王 知道 工程队 打听到 村民们 扩建了 哪座水坝.

Mr. W. **know** construction-team **find out** villagers rebuild **which dam**

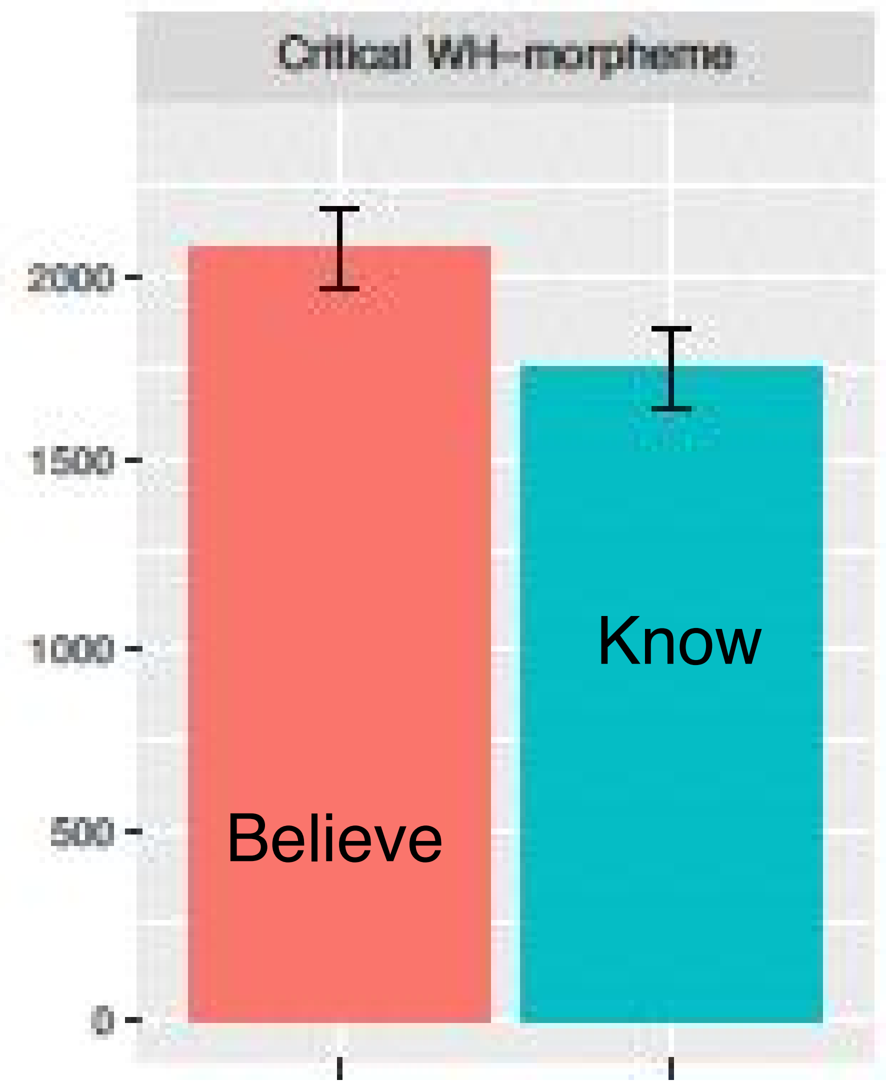
d.

小王 相信 工程队 打听到 村民们 扩建了 哪座水坝.

Mr. W. **believe** construction-team **find out** villagers rebuild **which dam**

Condition a&b, replicating Experiment 1

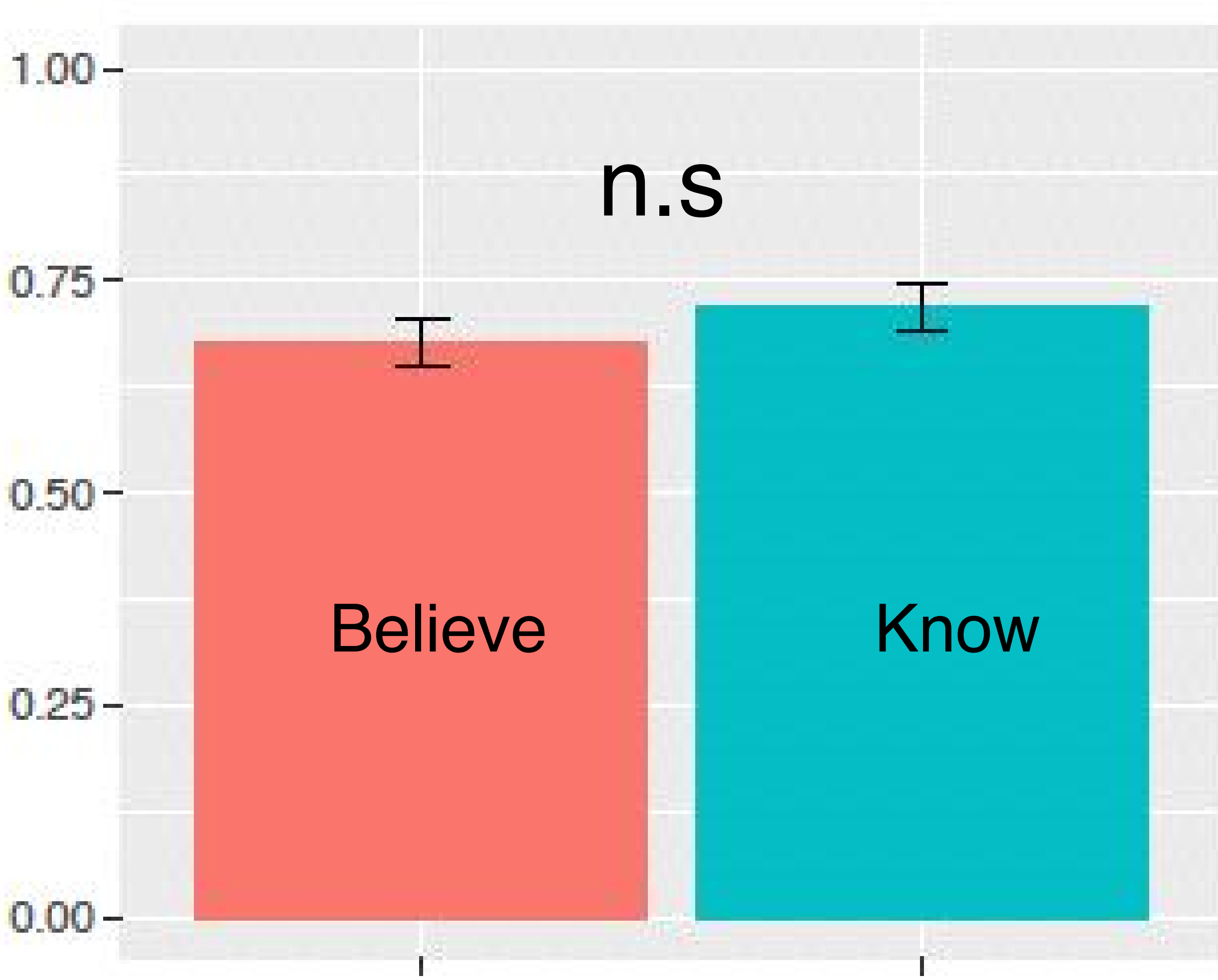
lower verbs: believe vs. know



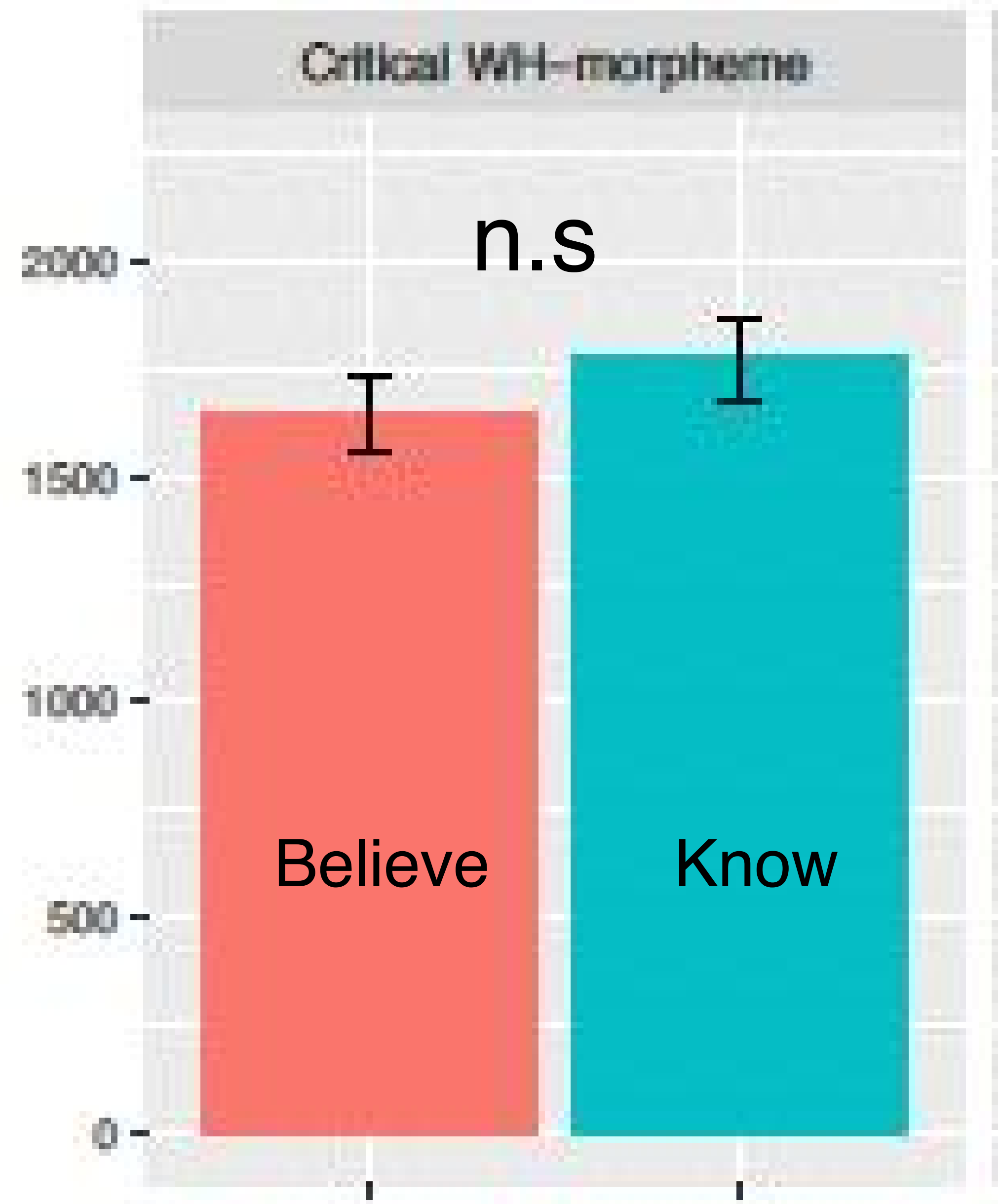
Acceptability judgment   Regression reading time

Condition c&d

Higher verbs: believe vs. know



Acceptability judgment



Regression reading time

Summary: the locality bias in parsing

- The observed locality bias is indeed driven by the parser's preference to associate the wh-in-situ expression with the local scope position.

- Ambiguity advantage is not a viable account of the data

How do we interpret wh-scope ambiguity?

$$\ldots V_{1}[_{CP1} \ldots V_{2}[_{CP2} \ldots WH]$$

If there is a locality bias when the parser establishes the scope dependency for the in-situ-wh, and if there is a direct mapping between parsing and interpretation, we should also expect a locality bias in interpretation.

Experiment 3

Truth value judgment task (subj =88, item =16)

Context:

At a recent archaeology conference, Emily said that her research team found evidence to prove that a famous ancient city was actually built by aliens. But she didn't release the name of the city.

"You will read a sentence below. Please indicate whether this sentence is consistent or inconsistent with the context above. "

Target sentence (ambiguous)

a. Emily announced her team discovered aliens built which city.

- ○ True
- ○ False

Target sentence (ambiguous)

a. Emily announced her team discovered aliens built which city.

low scope parse

"Emily announced her team **discovered which city** aliens built."

Target sentence (ambiguous)

a. Emily announced her team discovered aliens built which city.

low scope parse

- **True**

"Emily announced her team **discovered which city** aliens built."

Target sentence (ambiguous)

a. Emily announced her team discovered aliens built which city.

high scope parse

“Emily **announced** **which city** her team discovered aliens built.”

Target sentence (ambiguous)

a. Emily announced her team discovered aliens built which city.

high scope parse

- **False**

"Emily **announced which city** her team discovered aliens built."

Target sentence (ambiguous)

a. Emily announced her team discovered aliens built which city.

- **True**

  low scope

  "Emily announced her team discovered **which city** aliens built."

- **False**

  high scope

  "Emily announced **which city** her team discovered aliens built."

Target sentence (ambiguous)

b. Emily hid her team discovered aliens built which city.

- **False**

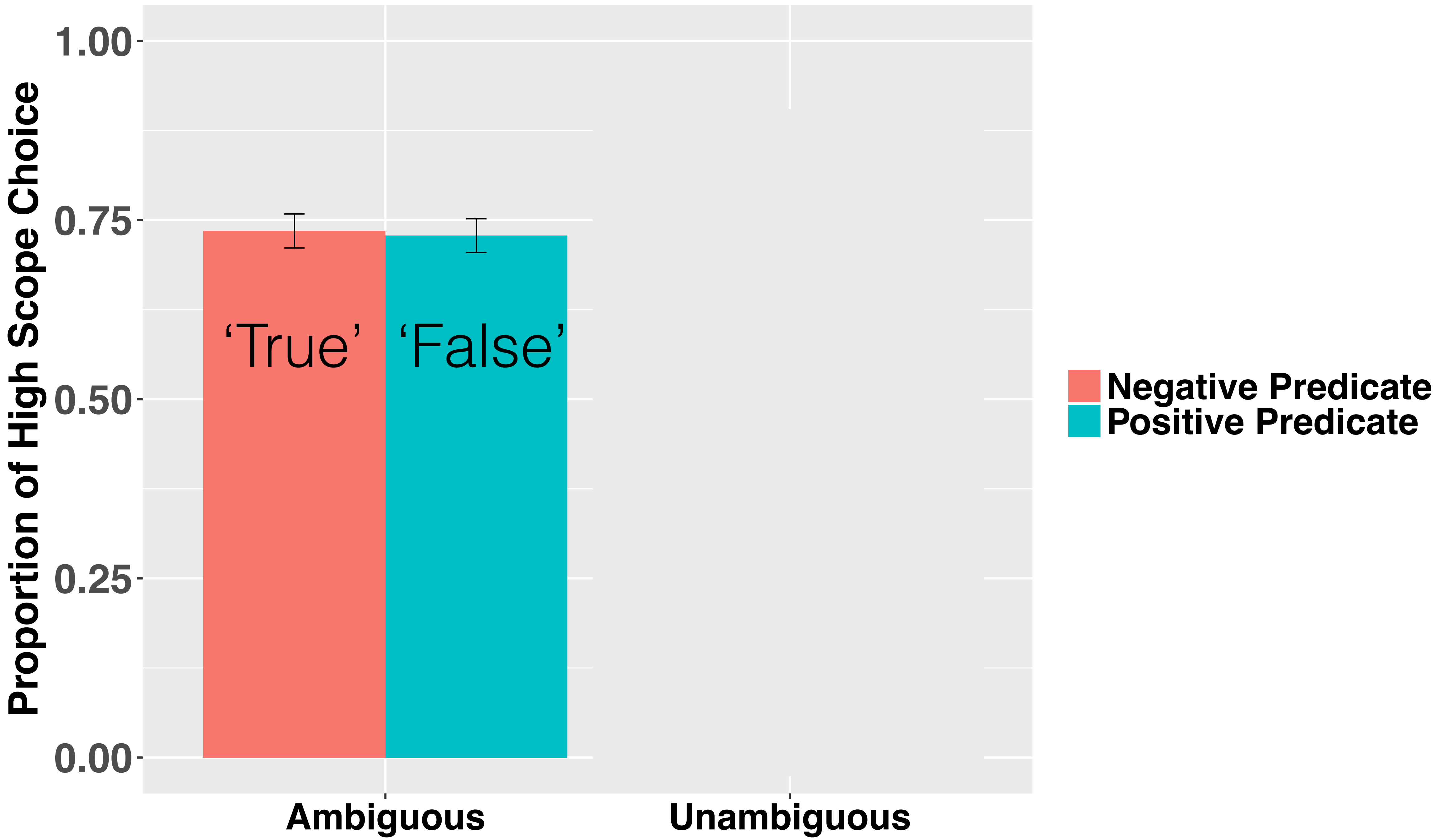
  low scope

  “Emily hid her team **discovered which city** aliens built.”

- **True**

  high scope

  “Emily **hid which city** her team discovered aliens built.”



Proportions of **high scope** choice

unambiguous control conditions:

$$\ldots V_1[_{CP1} \ldots V_2[_{CP2} \ldots WH]$$

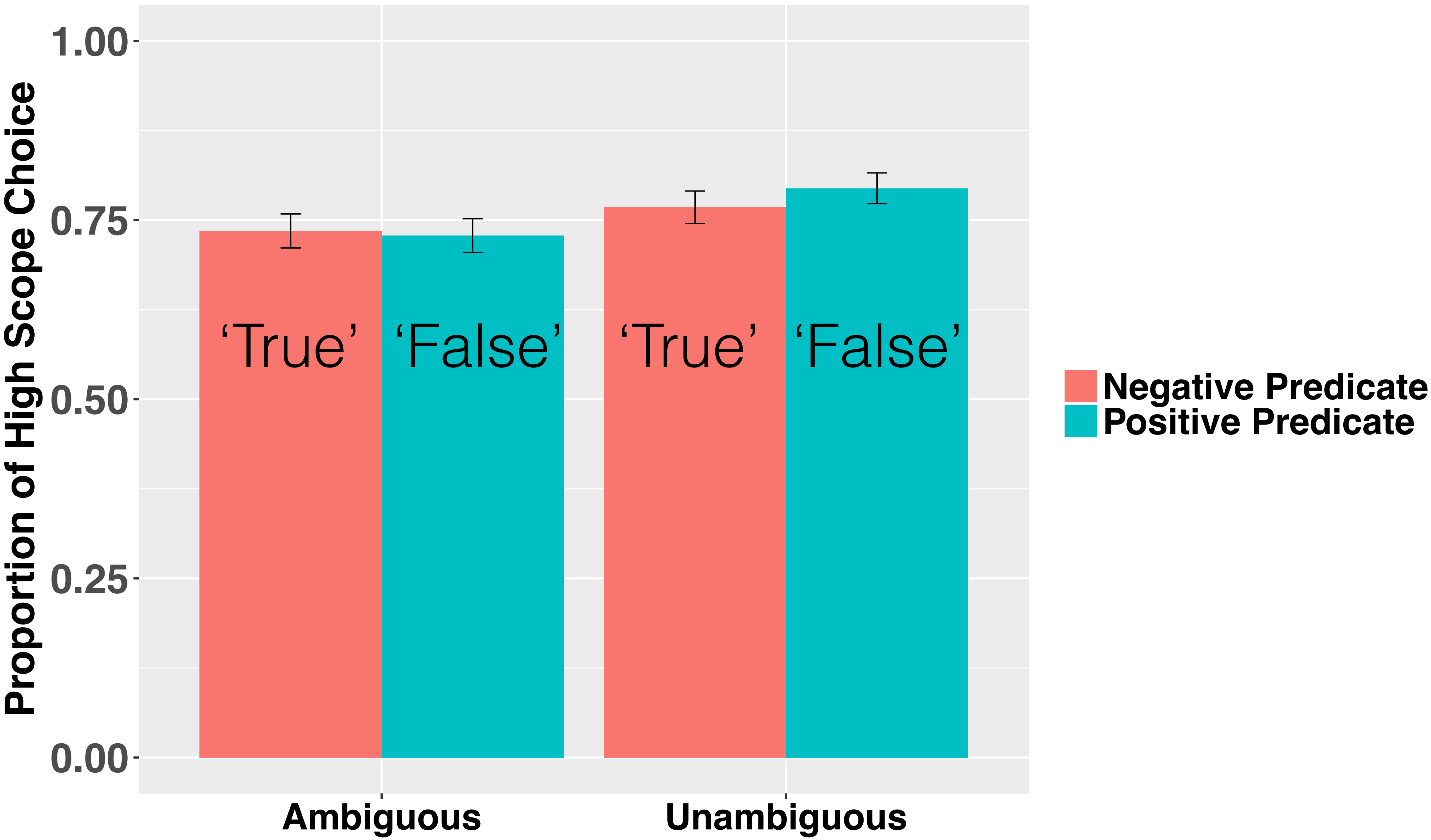
announce/hide **believe**

Context:

At a recent archaeology conference, Emily said that her research team found evidence to prove that a famous ancient city was actually built by aliens. But she didn’t release the name of the city.

Target:

Emily announced/hid her team believed aliens built which city.



Proportions of **high scope** choice

Experiment 4: reproducing the locality effect with acceptability judgments

- Exactly the same stimuli and procedure as Experiment 1
- But the task is a Yes/No acceptability judgment task
- Subj N=24

Experiment 4:

Context:

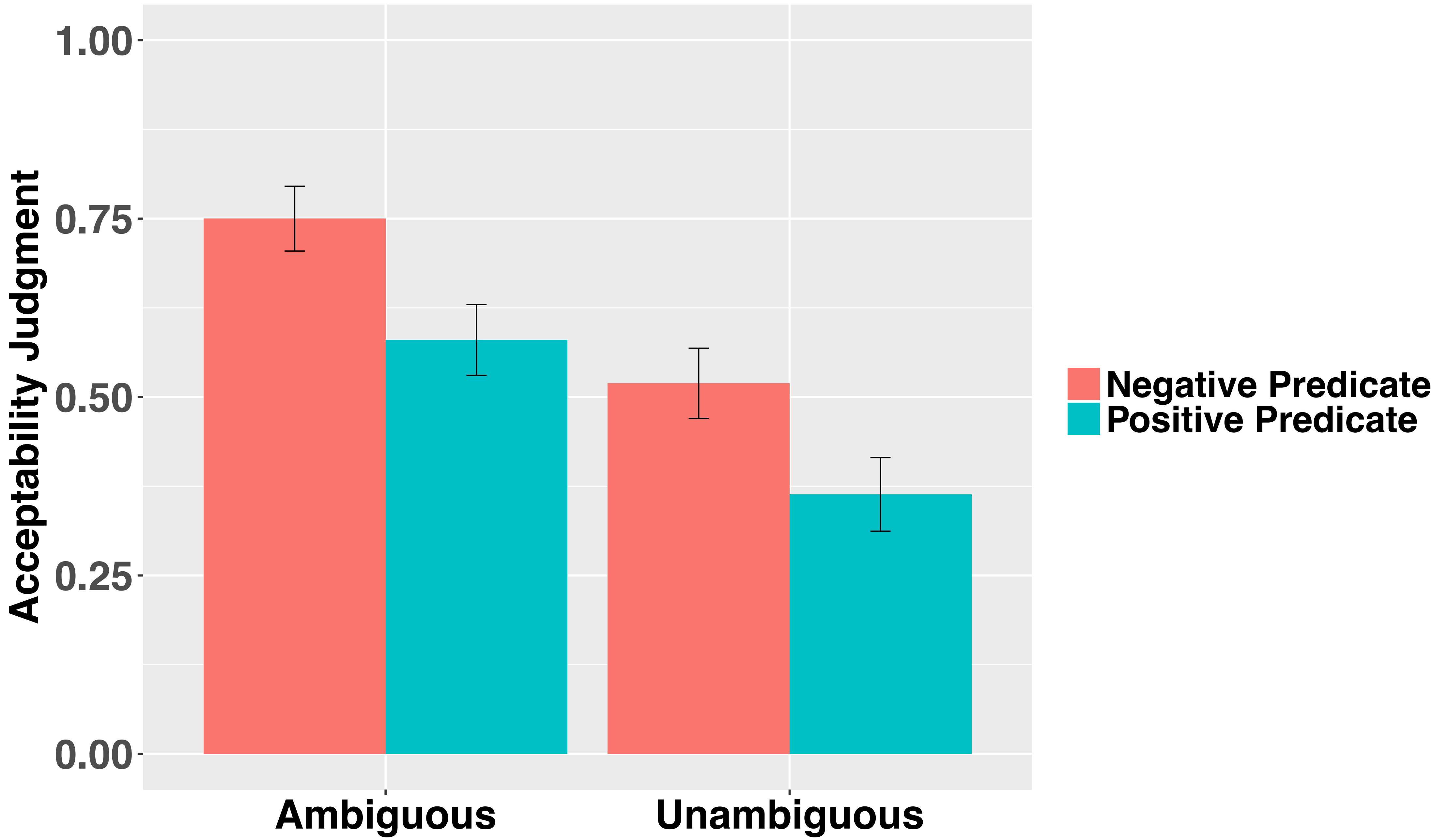
At a recent archaeology conference, Emily said that her research team found evidence to prove that a famous ancient city was actually built by aliens. But she didn't release the name of the city.

Target (acceptability judgment):

(Ambiguous): Emily announced/hid her team discovered aliens built which city.

(Unambiguous): Emily announced/hid her team believed aliens built which city.

- A locality bias in parsing would predict the unambiguously high-scope conditions, in which the local scope is blocked, should be judged less acceptable than the ambiguous conditions, in which the local scope dependency is available.



Proportions of “**yes**-acceptable” judgment

Summary and the puzzle

- The acceptability task replicates the locality bias
- Interpretations compatible with the high scope parse are preferred for ambiguous sentences

Where went wrong?

- The original linking hypothesis for a truth value judgment task

utterance

↓

parse tree

↓

semantics composed from the parse tree

↓

comparing the derived semantics with the context and make a T/F judgment

A new proposal

- Interpretation, as measured here by the truth value judgment task, is not uniquely determined by parsing

- We will model interpretation as the listener’s pragmatic inference about the relevant world states conditioned on the utterance they hear

- But we shouldn’t throw away parsing either. We can situate the parsing output within a Bayesian pragmatic inference model

Interpretation as Bayesian pragmatic inferences

$$P_L(w|u) = \frac{P_S(u|w) \times P_L(w)}{\sum_{w'} P_S(u|w') \times P_L(w')}$$

Interpretation as Bayesian pragmatic inferences

$$\underline{P_L(w|u)} = \frac{P_S(u|w) \times P_L(w)}{\sum_{w'} P_S(u|w') \times P_L(w')}$$

listeners carry out a probabilistic update about the relevant world states given they heard an utterance

Interpretation as Bayesian pragmatic inferences

The probability of a speaker producing the utterance given a relevant world state

$$P_L(w|u) = \frac{\underline{P_S(u|w)}\times P_L(w)}{\sum_{w'} P_S(u|w')\times P_L(w')}$$

Interpretation as Bayesian pragmatic inferences

The prior probability of the relevant world state

$$P_L(w|u) = \frac{P_S(u|w) \times \underline{P_L(w)}}{\sum_{w'} P_S(u|w') \times P_L(w')}$$

Rational speech act (RSA): the recursive reasoning between listeners and speakers (Goodman and Frank, 2016)

$$P_L(w|u) = \frac{P_S(u|w) \times P_L(w)}{\sum_{w'} P_S(u|w') \times P_L(w')}$$

The pragmatic listener

$$P_S(u|w) \propto exp(\alpha \times U_S(u;w))$$

The pragmatic speaker

$$U_S(u;w) = ln(L_0(w|u))$$

$$L_0(w|u) = \frac{\delta_{[u](w)} P(w)}{\sum_{w' \in W} \delta_{[u](w')} P(w')}$$

The literal listener

The recursive reasoning between listeners and speakers

Experiment 6 **Experiment 5**

$$P_L(w|u) = \frac{P_S(u|w) \times P_L(w)}{\sum_{w'} P_S(u|w') \times P_L(w')}$$

The pragmatic listener

$$P_S(u|w) \propto exp(\alpha \times U_S(u;w))$$

The pragmatic speaker

making predictions for production!

$$U_S(u;w) = ln(L_0(w|u))$$

$$L_0(w|u) = \frac{\delta_{[u](w)} P(w)}{\sum_{w' \in W} \delta_{[u](w')} P(w')}$$

The literal listener

A note about the literal listener:

1 or 0

Experiment 5

$$L_0(w|u) = \frac{\delta_{[\![u]\!](w)} P(w)}{\sum_{w' \in W} \delta_{[\![u]\!](w')} P(w')}$$

we can locate the parsing bias at the L_0 stage!

An outline for the rest of the talk

- Empirically estimate the priors for the relevant world state ***P(w)*** — Experiment 5

- **Compute the literal listener L_0, and make a qualitative prediction for the pragmatic speaker**

- Empirically estimate the production profile from a pragmatic **speaker** $P_S(u|w)$ — Experiment 6

- compute the pragmatic **listener** $P_L(w|u)$, and link the results to the truth value judgments

Defining the relevant “w” (world states) for the *positive predicate*

Emily announced her team discovered aliens built which city.

world states	E_1: Emily announced the name of the city under discovery	E_2: Emily announced their discovery
W_1	+	+
W_2	+	-
W_3	-	+
W_4	-	-

Defining the relevant "w" (world states) for the *negative predicate*

Emily hid her team discovered aliens built which city.

world states	E_1: Emily hid the name of the city under discovery	E_2: Emily hid their discovery
W_1	+	+
W_2	+	-
~~W_3~~	-	+
~~W_4~~	-	-

A summary of the relevant world states ("*w*") considered in our model

world states	Positive matrix predicate	Negative matrix predicate
w_1	Emily announced they discovered which city was built by aliens and she also announced the name of the city.	Emily hid the fact that they discovered which city was built by aliens and (necessarily) also hid the name of the city.
w_2	Emily announced they discovered which city was built by aliens but she did not announce the name of the city.	Emily did not hide the fact that they discovered which city was built by aliens but she hid the name of the city.

Experiment 5: estimating the priors for each world state (subject n=119; item n=16)

Dependent variable:

We calculated the preference proportion for each of the world states

A background context

Basic procedure

A force choice task between two situations that represent the two world states under consideration

What we get:

The prior probability of each world state, **p(w_1)** and **p(w_2)** under each type of predicate

world states	Positive matrix predicate	Negative matrix predicate
w_1	Emily announced they discovered which city was built by aliens and she also announced the name of the city.	Emily hid the fact that they discovered which city was built by aliens and (necessarily) also hid the name of the city.
w_2	Emily announced they discovered which city was built by aliens but she did not announce the name of the city.	Emily did not hide the fact that they discovered which city was built by aliens but she hid the name of the city.

Example trial for the w_1 and w_2 under the positive predicate:

Context

At a recent archaeology conference, Emily reported on work from her research team.

Question

"Which of the following situation is more likely to happen?"

"In her report, Emily said that her research team found evidence to prove that a famous ancient city was actually built by aliens. She also released the name of the city.

"In her report, Emily said that her research team found evidence to prove that a famous ancient city was actually built by aliens. But she didn't release the name of the city."

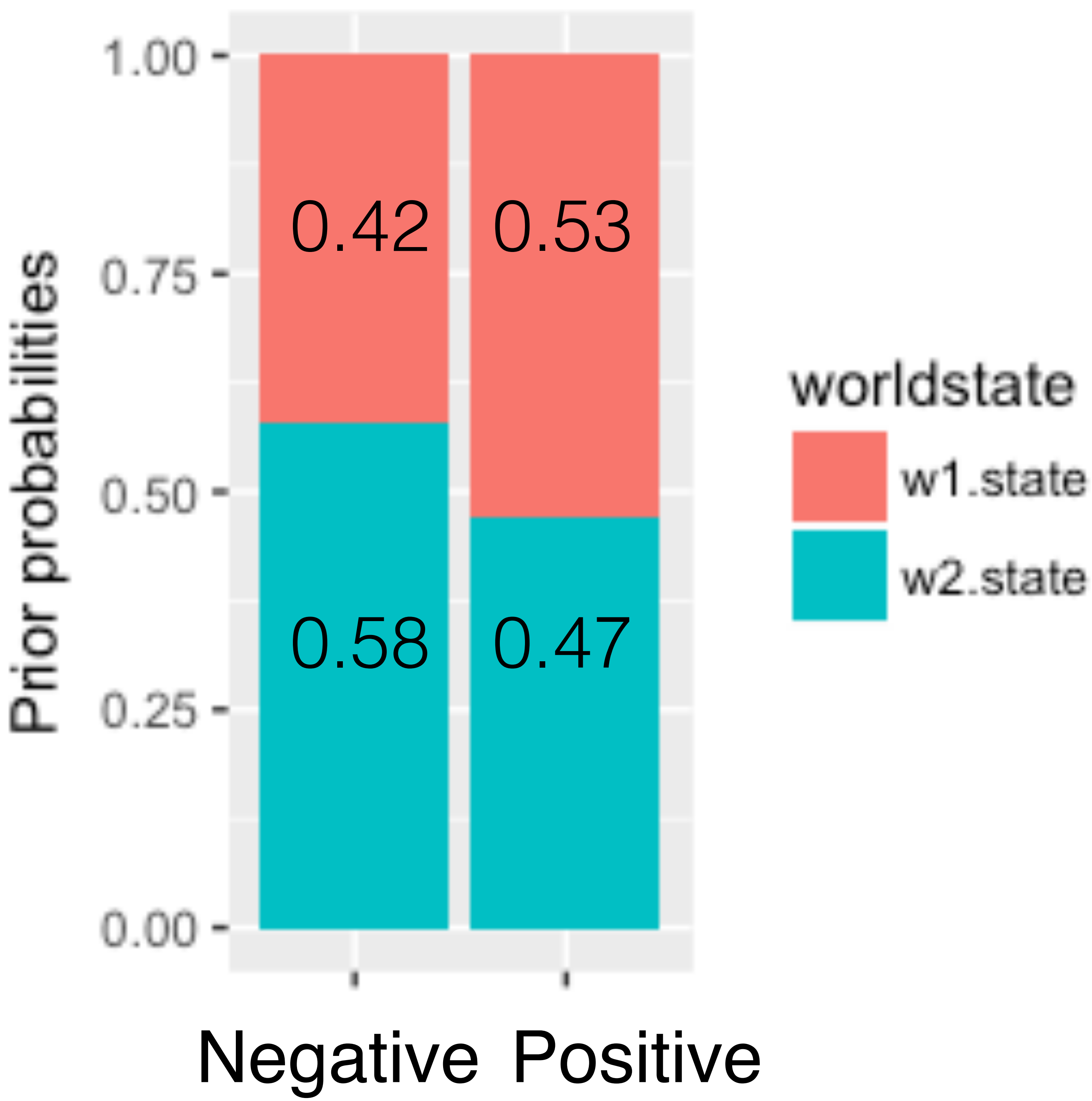
Example trial for the w_1 and w_2 under the negative predicate:

Context	At a recent archaeology conference, Emily reported on work from her research team.
Question	“Which of the following situation is more likely to happen?”

“Emily’s team actually have found evidence that an ancient city was built by aliens. But she completely concealed their discovery in her report.

“In her report, Emily said that her research team found evidence to prove that a famous ancient city was actually built by aliens. But she didn’t release the name of the city.”

The average prior probabilities of the world states



An outline for the rest of the talk

- Empirically estimate the priors for the relevant world state ***P(w)*** — Experiment 5
- **Compute the literal listener L_0, and make a qualitative prediction for the pragmatic speaker**
- Empirically estimate the production profile from a pragmatic **speaker $P_S(u|w)$** — Experiment 6
- compute the pragmatic **listener $P_L(w|u)$**, and link the results to the truth value judgments

The pragmatic inference of the listener

Experiment 6

Experiment 5

$$P_L(w|u) = \frac{P_S(u|w) \times P_L(w)}{\sum_{w'} P_S(u|w') \times P_L(w')}$$

Experiment 6: Empirically estimating the production bias (subject n=248, item n=16)

Dependent variable:

We calculated the proportion of the ambiguous wh-in-situ sentence form (as the ones used in the truth value judgment task)

Context that presents a particular world state

Basic procedure

Using the given fragments, participants were asked to produce sentences that are compatible with the given context

What we get:

The probability of a speaker producing the target wh-in-situ form given each world state: $\mathbf{p(u|w_1)}$ and $\mathbf{p(u|w_2)}$ under each predicate

world states	Positive matrix predicate	Negative matrix predicate
w_1	Emily announced they discovered which city was built by aliens and she also announced the name of the city.	Emily hid the fact that they discovered which city was built by aliens and (necessarily) also hid the name of the city.
w_2	Emily announced they discovered which city was built by aliens but she did not announce the name of the city.	Emily did not hide the fact that they discovered which city was built by aliens but she hid the name of the city.

Example trial: w_1 context (under the positive predicate)

Context	At a recent archaeology conference, Emily said that her research team found evidence to prove that a famous ancient city was actually built by aliens. She also released the name of the city.
Fragments for production	Emily **announced** / which city / built / her team discovered

Example trial: w_2 context (under the positive predicate)

Context

At a recent archaeology conference, Emily said that her research team found evidence to prove that a famous ancient city was actually built by aliens. But she didn't release the name of the city.

Fragments for production

Emily **announced**

which city

built

her team discovered

Example trial: w_1 context (under the negative predicate)

Context

Emily's research team found evidence to prove that a famous ancient city was actually built by aliens. But at a recent archaeology conference, she didn't mention this finding at all.

Fragments for production

Emily **hid**

which city

built

her team discovered

Example trial: w_2 context (under the negative predicate)

Context

At a recent archaeology conference, Emily said that her research team found evidence to prove that a famous ancient city was actually built by aliens. But she didn't release the name of the city.

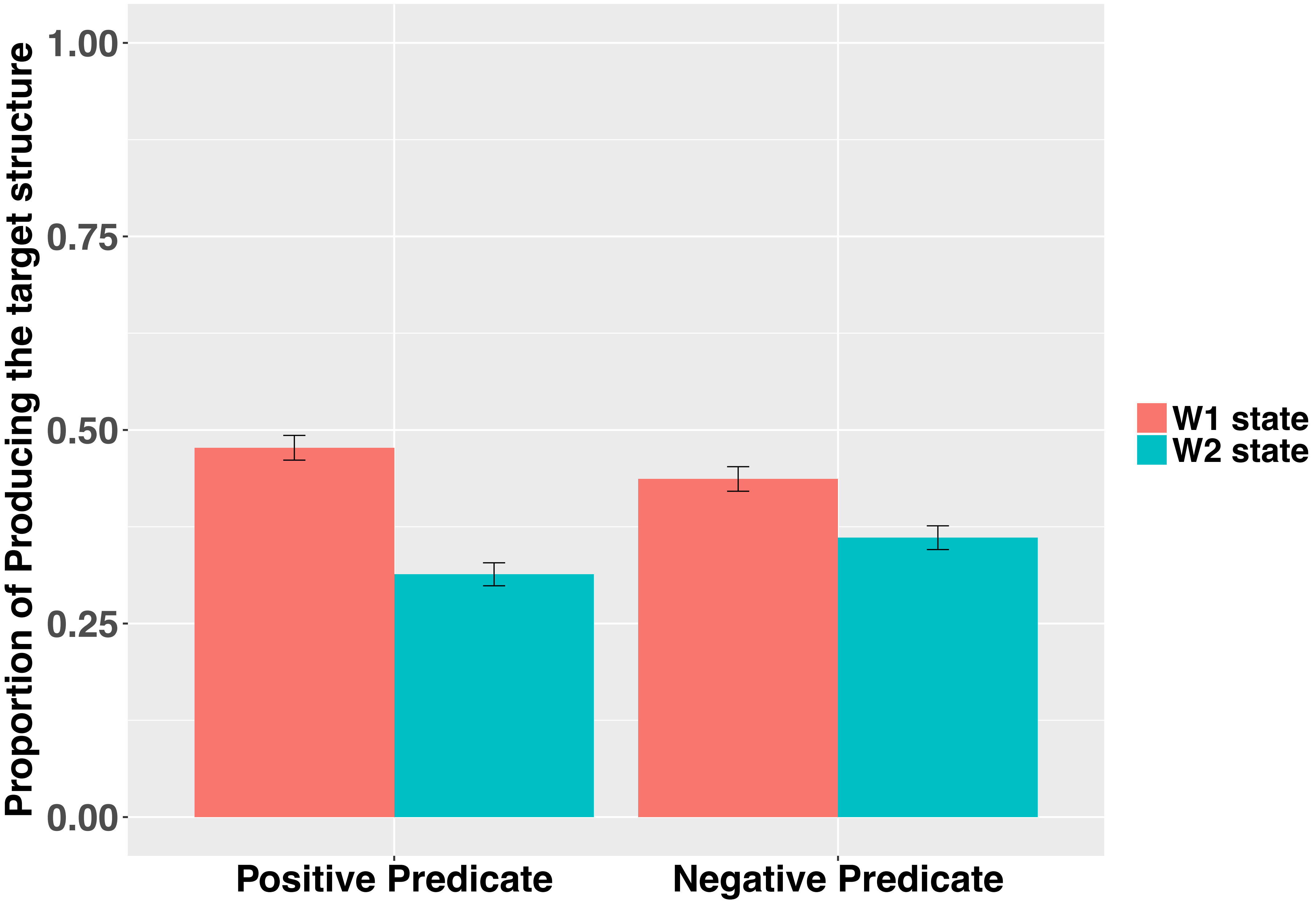
Fragments for production

Emily **hid**

which city

built

her team discovered



Proportion of Producing the target structure
1.00
0.75
0.50
0.25
0.00
Positive Predicate
Negative Predicate
W1 state
W2 state

An outline for the rest of the talk

- Empirically estimate the priors for the relevant world state ***P(w)*** — Experiment 5
- **Compute the literal listener L_0, and make a qualitative prediction for the pragmatic speaker**
- Empirically estimate the production profile from a pragmatic **speaker** $P_S(u|w)$ — Experiment 6
- compute the pragmatic **listener** $P_L(w|u)$, and link the results to the truth value judgments

The pragmatic inference of the listener

Experiment 6 Experiment 5

$$P_L(w|u) = \frac{P_S(u|w) \times P_L(w)}{\sum_{w'} P_S(u|w') \times P_L(w')}$$

Emily announced her team discovered aliens built which city.

world states	$P_s(u\|w)$	Priors	$P_L(w\|u)$
w_1: Emily announced they discovered which city was built by aliens and she also announced the name of the city.	0.48	0.53	**0.64**
w_2: Emily announced they discovered which city was built by aliens but she did not announce the name of the city.	0.31	0.47	**0.36**

$$P_{L_1}(w_1|u_{positive}) = \frac{P_{S1}(u|w_1) \times P_{L_1}(w_1)}{\sum_{w'} P_{S1}(u|w') \times P_{L_1}(w')} = \frac{0.48 \times 0.53}{0.48 \times 0.53 + 0.31 \times 0.47} = 0.64$$

$$P_{L_1}(w_2|u_{positive}) = \frac{P_{S1}(u|w_2) \times P_{L_1}(w_2)}{\sum_{w'} P_{S1}(u|w') \times P_{L_1}(w')} = \frac{0.31 \times 0.47}{0.48 \times 0.53 + 0.31 \times 0.47} = 0.36$$

Revisit the truth value judgment task in experiment 3

Context: At a recent archaeology conference, Emily said that her research team found evidence to prove that a famous ancient city was actually built by aliens. But she didn't release the name of the city.

Emily announced her team discovered aliens built which city.

world states	$P_s(u\|w)$	Priors	$P_L(w\|u)$
w_1: Emily announced they discovered which city was built by aliens and she also announced the name of the city.	0.48	0.53	**0.64**
w_2: Emily announced they discovered which city was built by aliens but she did not announce the name of the city.	0.31	0.47	**0.36**

triggers more "**False**" response under the given context

Emily hid her team discovered aliens built which city.

world states	$P_s(u\|w)$	Priors	$P_L(w\|u)$
w_1: Emily hid the fact that they discovered which city was built by aliens and (necessarily) also hid the name of the city.	0.44	0.42	**0.47**
w_2: Emily did not hide the fact that they discovered which city was built by aliens but she hid the name of the city.	0.36	0.58	**0.53**

$$P_L(w_1|u_{negative}) = \frac{P_S(u|w_1) \times P_L(w_1)}{\sum_{w'} P_S(u|w') \times P_L(w')}$$

$$= \frac{0.44 \times 0.42}{0.44 \times 0.42 + 0.36 \times 0.58}$$

$$= 0.47$$

$$P_L(w_2|u_{negative}) = \frac{P_S(u|w_2) \times P_L(w_2)}{\sum_{w'} P_S(u|w') \times P_L(w')}$$

$$= \frac{0.36 \times 0.58}{0.44 \times 0.42 + 0.36 \times 0.58}$$

$$= 0.53$$

Revisit the truth value judgment task in experiment 3

Context: At a recent archaeology conference, Emily said that her research team found evidence to prove that a famous ancient city was actually built by aliens. But she didn't release the name of the city.

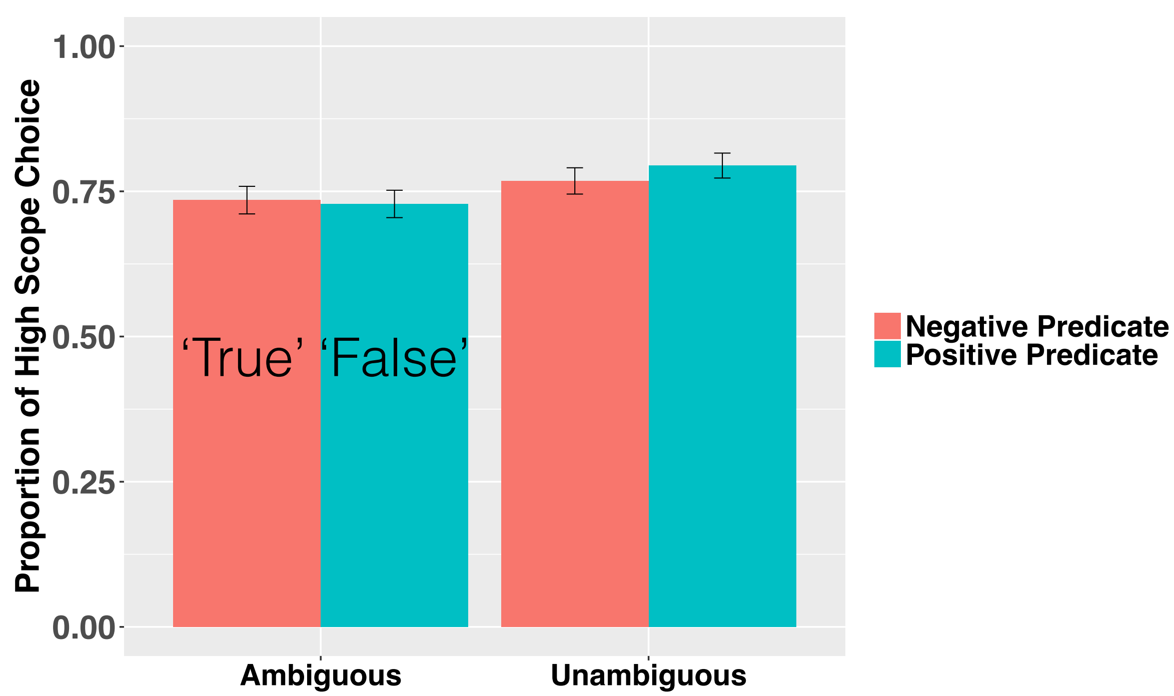
Emily hid her team discovered aliens built which city.

world states	$P_s(u\|w)$	Priors	$P_L(w\|u)$
w_1: Emily hid the fact that they discovered which city was built by aliens and (necessarily) also hid the name of the city.	0.44	0.42	**0.47**
w_2: Emily did not hide the fact that they discovered which city was built by aliens but she hid the name of the city.	0.36	0.58	**0.53**

triggers more "**True**" response under the given context

But 1…

- One remaining problem is that although the model prediction qualitatively matches the empirical results from the truth value judgments, it seems to under estimate, especially for utterances with negative matrix predicates



Empirical estimates from the truth value judgment task in Experiment 3

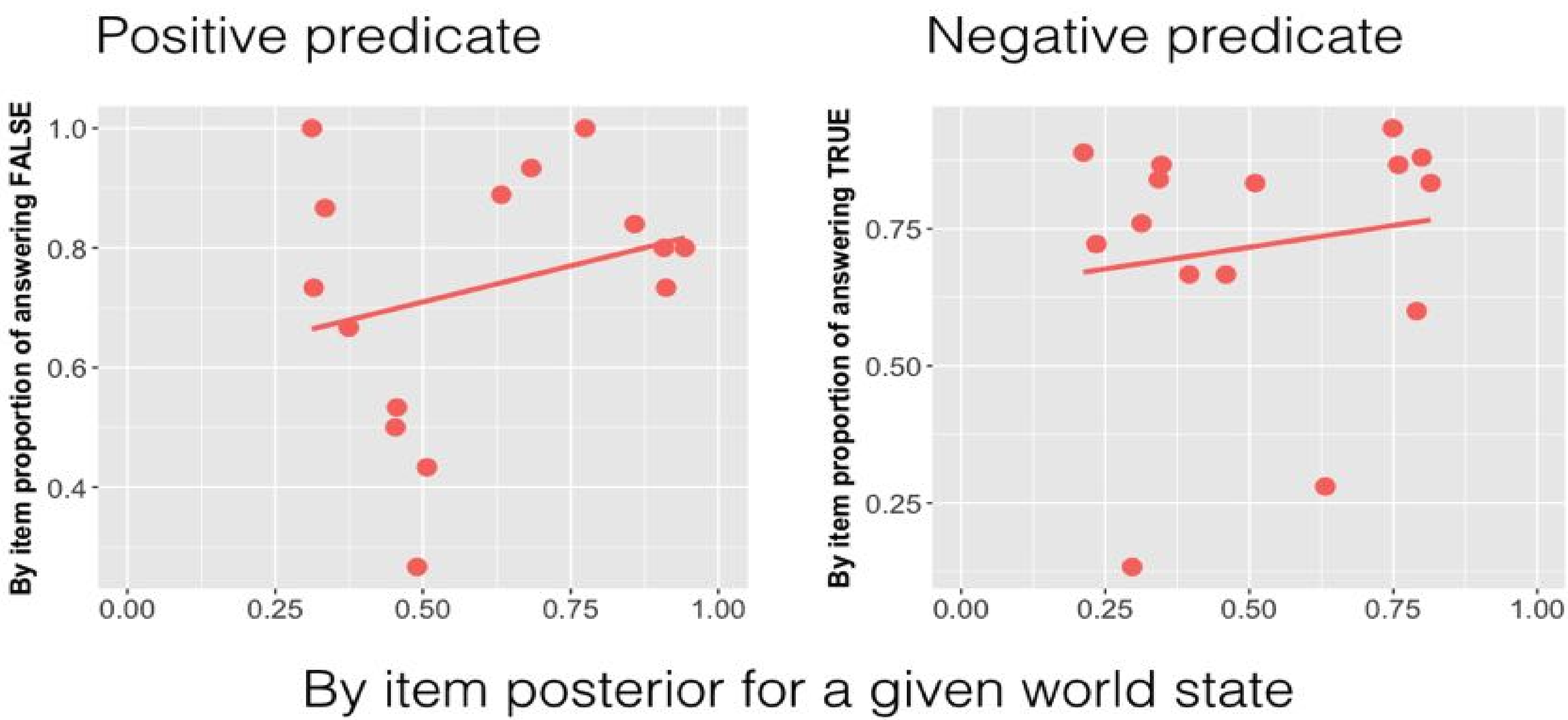
- A post-hoc thought:

  Comprehenders may have focused on applying their updated belief $\boldsymbol{P_L(w|u)}$ to answer a salient QUD raised by the context, instead of paying attention to all aspects of the context

  But it is an open question how to track QUDs in principled ways

But 2…

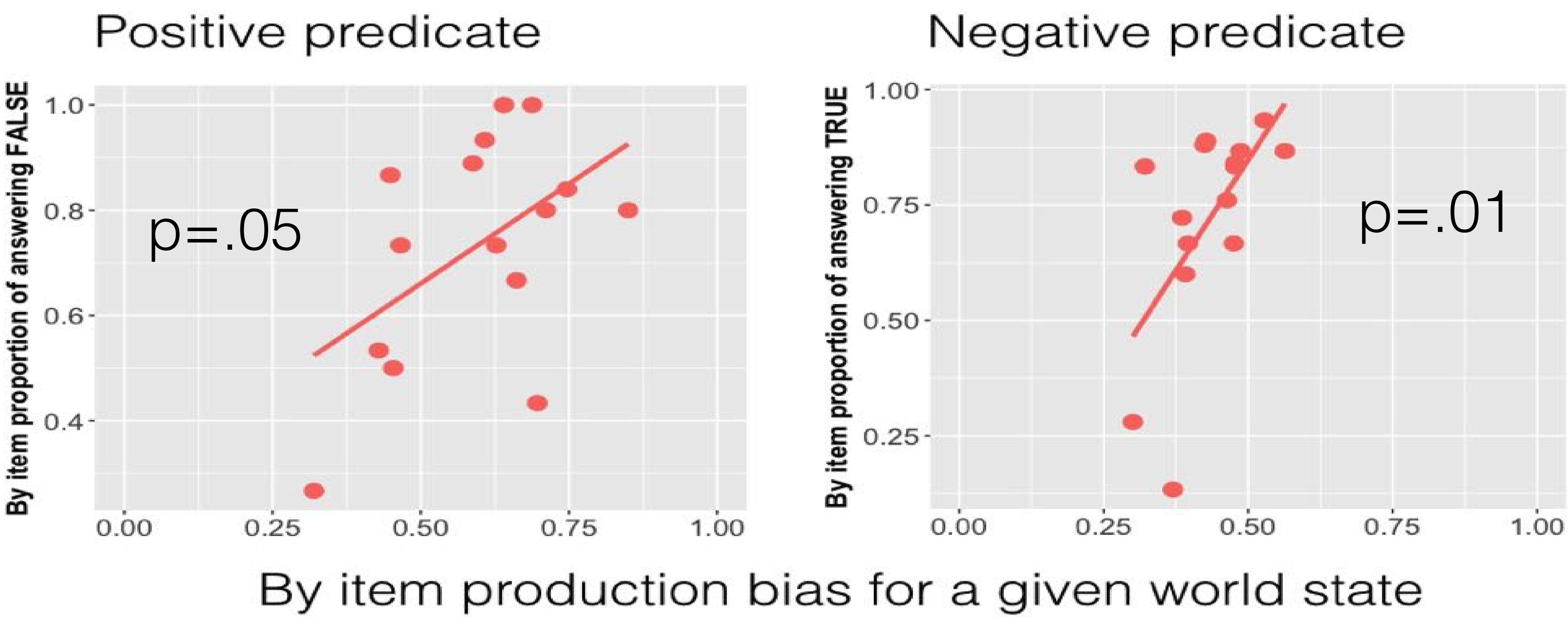
- The by-item fit based on both the prior and the production estimates is poor



$$P_L(w|u) \propto \frac{P_S(u|w) * P_L(w)}{\sum_{w'} P_S(u|w') * P_L(w')}$$

Not a good fit by item

- However, the by-item fit based on the production bias alone is reasonable



$$P_L(w|u) \propto \frac{P_S(u|w) * P_L(w)}{\sum_{w'} P_S(u|w') * P_L(w')}$$

reasonably good fit by item if assuming P(w1)=P(w2)=0.5

- This seems to suggest that comprehenders' pragmatic reasoning is more sensitive to the production alternatives for a given message (e.g. what the speaker could have said) than the prior likelihood of a message

- But the lack of sensitivity to priors may be due to the fact that the current context scenarios are in general ad-hoc ones

What we have done

- When interpretation is modeled as a listener's probabilistic belief update about the relevant world states, we can derive the truth value judgment results

- The probabilistic update can be modeled as Bayesian pragmatic inferences

- There is no real misalignment between parsing and interpretation in the end, since parsing outcomes are integrated into the pragmatic reasoning model

Experiment 6 **Experiment 5**

$$P_L(w|u) = \frac{P_S(u|w) \times P_L(w)}{\sum_{w'} P_S(u|w') \times P_L(w')}$$

The pragmatic listener

$$P_S(u|w) \propto exp(\alpha \times U_S(u;w))$$

The pragmatic speaker

$$U_S(u;w) = ln(L_0(w|u))$$

making predictions for production!

$$L_0(w|u) = \frac{\delta_{[u](w)} P(w)}{\sum_{w' \in W} \delta_{[u](w')} P(w')}$$

The literal listener

we can locate the parsing bias here

- At the theoretical level, we introduce a new analytical possibility to account for possible "misalignments" between parsing and interpretation

- The current approach has the added benefit to potentially link language production and comprehension within the same processing architecture

Empirical and methodological implications:

- Empirically, we showed a kind of ‘intervention’ effect due to parsing bias: intervening CP positions that do not host wh-scope lead to substantial processing difficulty

- Such difficulty could be rescued to some degree by an anticipatory cue that signals the grammatical scope position

Acceptability judgments task

- Acceptability judgments, as already known in the literature, are sensitive to not just grammatical well-formedness but also parsing complexity

Truth value judgment task

utterance

↓

parse tree

↓

semantics composed from the parse tree

↓

comparing **the derived semantics** with the context and make a T/F judgment

---

utterance

↓

parse tree

↓

semantics composed from the parse tree

belief update about the possible messages/world state the speaker delivers

↓

comparing **the updated belief** with the context and make a T/F judgment

Suiping Wang
South China
Normal University

Greg Kobele
Leipzig University

Zhewei Dai
Alma College

Thank you!

The literal listener

$$L_0(w|u) = \frac{\delta_{[\![u]\!](w)} P(w)}{\sum_{w' \in W} \delta_{[\![u]\!](w')} P(w')}$$

The literal listener step is where pragmatics inferences get connected to compositional semantics

1 or 0 empirically estimated in Expt 5

$$L_0(w|u) = \frac{\delta_{[\![u]\!](w)} P(w)}{\sum_{w' \in W} \delta_{[\![u]\!](w')} P(w')}$$

Adding the parsing bias at *L₀*

$$L_0(w)$$

$$= L_0(w|u_h) \times P(u_h) + L_0(w|u_l) \times P(u_l)$$

$$= \frac{\delta_{[\![u_h]\!](w)} P(w)}{\sum_{w'} \delta_{[\![u_h]\!](w')} P(w')} \times P(u_h) + \frac{\delta_{[\![u_l]\!](w)} P(w)}{\sum_{w'} \delta_{[\![u_l]\!](w')} P(w')} \times P(u_l)$$

For the **positive** predicate:

Emily announced her team discovered aliens built which city.

High: Emily announced **which city** her team discovered aliens built.

Low: Emily announced her team discovered **which city** aliens built.

world states	Positive matrix predicate	$P(w)$	$\delta_{[u_h](w)}$	$\delta_{[u_l](w)}$
w_1	Emily announced they discovered which city was built by aliens and she also announced the name of the city.	0.53	1	1
w_2	Emily announced they discovered which city was built by aliens but she did not announce the name of the city.	0.47	0	1

For the **positive** predicate:

$$L_0(w_1)$$

$$= \frac{\delta_{[u_h](w_1)}P(w_1)}{\delta_{[u_h](w_1)}P(w_1) + \delta_{[u_h](w_2)}P(w_2)} \times P(u_h) + \frac{\delta_{[u_l](w_1)}P(w_1)}{\delta_{[u_l](w_1)}P(w_1) + \delta_{[u_l](w_2)}P(w_2)} \times P(u_l)$$

$$= \frac{1 \times 0.53}{1 \times 0.53 + 0 \times 0.47} \times \underline{0.5} + \frac{1 \times 0.53}{1 \times 0.53 + 1 \times 0.47} \times \underline{0.5}$$

$$= 1 \times 0.5 + 0.53 \times 0.5$$

$$= 0.765$$

between **0.53** and **0.765** if assuming
0<p(u_h)<0.5 and 0.5<p(u_L)<1

$$L_0(w_2)$$

$$= \frac{\delta_{[u_h](w_2)}P(w_2)}{\delta_{[u_h](w_1)}P(w_1) + \delta_{[u_h](w_2)}P(w_2)} \times P(u_h) + \frac{\delta_{[u_l](w_2)}P(w_2)}{\delta_{[u_l](w_1)}P(w_1) + \delta_{[u_l](w_2)}P(w_2)} \times P(u_l)$$

$$= \frac{0 \times 0.47}{1 \times 0.53 + 0 \times 0.47} \times \underline{0.5} + \frac{1 \times 0.47}{1 \times 0.53 + 1 \times 0.47} \times \underline{0.5}$$

$$= 0 + 0.47 \times 0.5$$

$$= 0.235$$

For the **negative** predicate:

Emily hid her team discovered aliens built which city.

High: Emily hid **which city** her team discovered aliens built.

Low: Emily hid her team discovered **which city** aliens built.

world states	Negative matrix predicate	$P(w)$	$\delta_{[u_h](w)}$	$\delta_{[u_l](w)}$
w_1	Emily hid the fact that they discovered which city was built by aliens and (necessarily) also hid the name of the city.	0.42	1	1
w_2	Emily did not hide the fact that they discovered which city was built by aliens but she hid the name of the city.	0.58	1	0

For the **negative** predicate:

$$L_0(w_1)$$

$$= \frac{\delta_{[\![u_h]\!](w_1)} P(w_1)}{\delta_{[\![u_h]\!](w_1)} P(w_1) + \delta_{[\![u_h]\!](w_2)} P(w_2)} \times P(u_h) + \frac{\delta_{[\![u_l]\!](w_1)} P(w_1)}{\delta_{[\![u_l]\!](w_1)} P(w_1) + \delta_{[\![u_l]\!](w_2)} P(w_2)} \times P(u_l)$$

$$= \frac{1 \times 0.42}{1 \times 0.42 + 1 \times 0.58} \times 0.5 + \frac{1 \times 0.42}{1 \times 0.42 + 0 \times 0.58} \times 0.5$$

$$= 0.42 \times 0.5 + 1 \times 0.5$$

$$= 0.71$$

between **0.71** and **1** if assuming
$0<p(u_h)<0.5$ and $0.5<p(u_L)<1$

$$L_0(w_2)$$

$$= \frac{\delta_{[\![u_h]\!](w_2)} P(w_2)}{\delta_{[\![u_h]\!](w_1)} P(w_1) + \delta_{[\![u_h]\!](w_2)} P(w_2)} \times P(u_h) + \frac{\delta_{[\![u_l]\!](w_2)} P(w_2)}{\delta_{[\![u_l]\!](w_1)} P(w_1) + \delta_{[\![u_l]\!](w_2)} P(w_2)} \times P(u_l)$$

$$= \frac{1 \times 0.58}{1 \times 0.42 + 1 \times 0.58} \times 0.5 + \frac{0 \times 0.58}{1 \times 0.42 + 0 \times 0.58} \times 0.5$$

$$= 0.58 \times 0.5 + 0$$

$$= 0.29$$

The recursive reasoning between listeners and speakers

$$P_L(w|u) = \frac{P_S(u|w) \times P_L(w)}{\sum_{w'} P_S(u|w') \times P_L(w')}$$

The pragmatic listener

$$P_S(u|w) \propto exp(\alpha \times U_S(u;w))$$

The pragmatic speaker

making predictions for production!

$$U_S(u;w) = ln(L_0(w|u))$$

$$L_0(w|u) = \frac{\delta_{[u](w)} P(w)}{\sum_{w' \in W} \delta_{[u](w')} P(w')}$$

The literal listener

The pragmatic speaker

$$P_S(u|w) \propto exp(\alpha \times \underline{U_S(u;w)})$$

$$U_S(u;w) = ln(\underline{L_0(w|u)})$$

The pragmatic speaker

$$P_S(u|w) \propto exp(\alpha \times U_S(u;w))$$

$$U_S(u;w) = ln(L_0(w|u))$$

A qualitative prediction:

Since: $P_{L0}(w_1|u) > P_{L0}(w_2|u)$

Therefore: $P_s(u|w_1) > P_s(u|w_2)$